

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

Om nucleic - nucleic search, using sw model

Run on: September 25, 2002, 06:26:33 ; Search time 2465.47 Seconds
 (without alignments)
 1091.764 Million cell updates/sec

Title: US-09-762-249-13

Perfect score: 1295

Sequence: 1 atccagcgcacgtggat.....aaaaaaaaaaaaaaaaaaa 1295

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_ss:*

12: gb_sy:*

13: gb_uni:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_in:*

18: em_mu:*

19: em_nu:*

20: em_on:*

21: em_or:*

22: em_ov:*

23: em_pet:*

24: em_ph:*

25: em_pl:*

26: em_lo:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_hhg_hum:*

31: em_hhg_inv:*

32: em_htgo_inv:*

33: em_htgo_other:*

RESULT 1

HSA012506 HSA012506 LOCUS Homo sapiens mRNA activated in tumor suppression, clone TSA21 DEFINITION Homo sapiens mRNA activated in tumor suppression, clone TSA21 extended.

ACCESSION AJ012506

VERSION AJ012506.1 GI:5441365

KEYWORDS tumor suppression.

SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 1296)

AUTHORS Roperech,J.P., Lehrone,F., Prieur,S., Plouffe,L., Israeli,D., Tuynier,M., Neimani,M., Pastureaud,P., Gendron,M.C., Dausset,J., Oren,M., Amson,R.B. and Telerman,A.

TITLE SIAH-1 promotes apoptosis and tumor suppression through a network involving the regulation of protein folding, unfolding, and trafficking: identification of common effectors with p53 and p21(Waf1)

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (14), 8070-8073 (1999)

REFERENCE 2 (bases 1 to 1296)

ALIGMENTS

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	DB ID	Description
1	3595312			

Group. Further information can be found at

<http://www.sanger.ac.uk/HPG/CRC6>

RPI-83M4 is from the library RPCI-1 constructed by the group of

Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: PCYPAC2

IMPORTANT: This sequence is not the entire insert of clone RPI-83M4. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone RPI-83M4 is at 1 in this sequence. The

true left

end of clone RPI-91124 is at 104129 in this sequence.

FEATURES

source

```
1..104228
/organism="Homo sapiens"
/db_xref="taxon_9606"
/chromosome="6"
/clone="RPI-83M4"
/clone_id="RPCI-1"
1302..1605
/note="AlusX repeat: matches 1..303 of consensus"
2223..2533
/note="L2 repeat: matches 1866..2190 of consensus"
repeat_region
/note="MIR repeat: matches 202..247 of consensus"
repeat_region
2581..2634
/note="L2 repeat: matches 2693..2747 of consensus"
repeat_region
2637..2861
/note="MIR repeat: matches 1..262 of consensus"
repeat_region
3966..4024
/note="AluJ/FLAM repeat: matches 35..77 of consensus"
repeat_region
4613..4720
/note="27 copies 4 mer tctt 92% conserved"
4759..5035
/note="AlusP repeat: matches 1..278 of consensus"
repeat_region
5613..5697
/note="MIR repeat: matches 65..150 of consensus"
repeat_region
5746..6018
/note="LTR16A repeat: matches 169..450 of consensus"
6681..6805
/note="ME5A repeat: matches 61..185 of consensus"
repeat_region
7197..7476
/note="2 copies 140 mer 85% conserved"
repeat_region
7344..7704
/note="L2 repeat: matches 2355..2530 of consensus"
7835..7962
/note="FLAMC repeat: matches 2609..2746 of consensus"
repeat_region
8707..8977
/note="AlusX repeat: matches 1..268 of consensus"
repeat_region
10041..10339
/note="AlusX repeat: matches 1..301 of consensus"
11563..12086
/note="FLAMC repeat: matches 1..124 of consensus"
repeat_region
12102..12143
/note="AluS repeat: matches 98..142 of consensus"
repeat_region
12144..12456
/note="AlusX repeat: matches 1..304 of consensus"
repeat_region
12457..12617
/note="AluO repeat: matches 142..282 of consensus"
repeat_region
12648..12786
/note="L2 repeat: matches 2595..2745 of consensus"
repeat_region
13356..13668
/note="AlusQ repeat: matches 1..310 of consensus"
repeat_region
13845..14157
/note="AlusP repeat: matches 1..313 of consensus"
repeat_region
14484..14623
/note="MIR repeat: matches 70..212 of consensus"
repeat_region
15168..15497
/note="AlusX repeat: matches 1..311 of consensus"
repeat_region
15498..15557
/note="30 copies 2 mer cc 68% conserved"
repeat_region
15931..16276
/note="trigger4(Zombi) repeat: matches 2383..2731 of
consensus"
```

repeat_region 16277..16587
/note="AluJb repeat: matches 6..312 of consensus"
repeat_region 16588..18382
/note="trigger4(Zombi) repeat: matches 627..2383 of

consensus"
16997..17123
/note="Sequence from overlapping clone ba3B11 (AL512354).
Assembly confirmed by restriction digest."

repeat_region 18383..18643
/note="AluJb repeat: matches 7..275 of consensus"
repeat_region 18644..19263
/note="Trigger4 (Zombi) repeat: matches 1..527 of consensus"
repeat_region 19240..21242
/note="MLTH repeat: matches 70..429 of consensus"
repeat_region 22808..23017
/note="MLTH repeat: matches 154..258 of consensus"
repeat_region 23035..23336
/note="AluJb repeat: matches 1..309 of consensus"
repeat_region 23523..23658
/note="AluJb repeat: matches 1..139 of consensus"
repeat_region 23659..23917
/note="AluJb repeat: matches 70..285 of consensus"
repeat_region 24084..24278
/note="AlusX repeat: matches 1..195 of consensus"
repeat_region 25005..25087
/note="MIR repeat: matches 1..219 of consensus"
repeat_region 26139..26440
/note="AlusG repeat: matches 1..299 of consensus"
repeat_region 26889..27041
/note="MIR repeat: matches 78..252 of consensus"
repeat_region 27623..28017
/note="AlusX repeat: matches 1..307 of consensus"
repeat_region 29985..30269
/note="THE1B repeat: matches 1..364 of consensus"
repeat_region 30313..30457
/note="MER5A repeat: matches 2..189 of consensus"
repeat_region 32032..32933
/note="AlusX repeat: matches 1..354 of consensus"
repeat_region 33301..33342
/note="MIR repeat: matches 1..307 of consensus"
repeat_region 34707..34752
/note="MER5A repeat: matches 2..189 of consensus"
repeat_region 34811..33100
/note="LPA7 repeat: matches 5426..6134 of consensus"
repeat_region 34992..35031
/note="10 copies 4 mer aaat 85% conserved"
misc_feature 35362..36323
/note="CPG Island"
repeat_region 37387..37697
/note="AlusX repeat: matches 1..311 of consensus"
repeat_region 37710..37993
/note="AluO repeat: matches 1..286 of consensus"
repeat_region 38849..39089
/note="MIR repeat: matches 2..262 of consensus"
repeat_region 40658..40945
/note="AlusX repeat: matches 1..288 of consensus"
repeat_region 41221..41256
/note="18 copies 2 mer ac 100% conserved"
repeat_region 41267..41347
/note="L2 repeat: matches 1806..1889 of consensus"
repeat_region 41323..41464
/note="L2 repeat: matches 2007..2143 of consensus"
repeat_region 41487..41557
/note="LMB2 repeat: matches 6099..6167 of consensus"
repeat_region 41558..41965
/note="MSIA repeat: matches 1..426 of consensus"

Db 448 ATCCAGCGCAGCTAGAGATCACTGGAGGACCAACTGAGGGATG 507
 Qy 61 ttcgagcaggtaatggggactgtttccggagaacttgcgtggccacgtgaaaggcg 120
 Db 508 CTGGAGAGCGGGAGGCCATCACTTCATCTGGAGATTATTCAGTTCAAGATACT 567
 Qy 121 cggcgcgcctcaacagatcgagacgcggccaaactgtggcgctggagggcg 180
 Db 568 AGGCAAAGCTCTCATGAGTCAGTCAGGCCAAAGACATCATGAAGCTGGAGACCG 627
 Qy 181 atccggagacgtacacagactcttcgtcagatggcggtgtggagaacggcggc 240
 Db 628 ATCCAGAGGCCTCACCGAGATCTCATGATGGCATGGCTGAGCTAGGGTGA 687
 Qy 241 accctgaacgtcatcgactcaacggtacaataaaacgggtcgatcacccggcaggc 300
 Db 688 ATGGTCAACAACATTCAGAGAACGACGTGGTACCTCGTAGATACGGAGACGCCA 747
 Qy 301 gcgccgtggggaaaggccgtggactcggaggaaacccctgcgggaccctctgtgc 360
 Db 748 GAAGAGACTAAAGAACCCATCAATAACCAGAGCAAGGCCAGAGGGTGCTCTGNGT 807
 Qy 361 ttctgtgtgcctc 375
 Db 808 CTCGGCGAGCTGG 822

RESULT 10
 ARR11444
 LOCUS ARR11444
 DEFINITION Sequence 13 from patent US 6127149. 486 bp DNA linear PAT 14-FEB-2001
 ACCESSION ARR11444.1
 VERSION G1:12828292
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 486)
 AUTHORS Hirai,Y., Koshida,S and Oka,Y.
 MODIFIED epimorphin
 JOURNAL Patent: US 6127149-A 13-03-OCT-2000;
 FEATURES source
 /organism="unknown"
 /location/Qualifiers 1..486
 BASE COUNT 152 a 110 c 139 g 85 t.
 ORIGIN

'very Match 9.6%; score 124.8; DB 6; length 486;
 best Local Similarity 60.7%; Pred. No. 3.9e-10; Mismatches 204; Conservative 0; Indels 0; Gaps 0;

Qy 1 atcccgccgcacgtggatcatggccaaaggaaatctcgccggcggccatgtggatcg 60
 Db 133 ATCCAGCGCAGCTGGAGATCACTGGAGGACCAACTGAGGGATG 192
 Qy 61 ttccatggatggatggggactgtgtttccggaaacttgcgtggccacgtgaaaggcg 120
 Db 193 CTGGAGAGCGGGAGGCCATCTCATCTGGATATTATTCAGTTCAAACT 252
 Qy 121 cggcgcgcctcaacagatcgagacgcggccaaactgtggcgctggagggcg 180
 Db 253 AGGCAAAGCTCTCATGAGTCAGTCAGGCCAAAGACATCATGAAGCTGGAGACCG 312
 Qy 181 atccggacatcaacagacttttgcataatggcggtgtggatggaaaggccgc 240
 Db 313 ATCCAGCGCAGCTGGAGATCACTGGAGGACCAACTGAGGGATG 372
 Qy 241 accctgaacgtcatcgactcaacggtacaaggcggtcgactacccggcaggc 300
 Db 373 ATGGTCAACAACATTCAGAGAACGACGTGGTACCTCGTAGATACGGAGACATGGCAAG 432

Qy 301 gecgaggcgccggaggccgtgcgttgcggatggggaaag 336
 Db 433 GAAGAGCGAAGGCCATCAATACTGGAGCAAG 468

RESULT 11
 EL2728
 LOCUS E12728
 DEFINITION DNA encoding mouse Epimorphin-derived peptide.
 ACCESSION E12728
 VERSION E12728.1 GI:3251560
 KEYWORDS JP 1397065-885-A/7.
 SOURCE unidentified.
 ORGANISM unidentified.
 unclassified.
 REFERENCE 1 (bases 1 to 486)
 AUTHORS Koshida,S., Oka,Y. and Hirai,Y.
 TITLE TAILORED DERIVATIVE OF EPIMORPHIN
 JOURNAL PATENT: JP 1997065885-A 7 11-MAR-1997;
 COMMENT SUMITOMO ELECTRIC IND LTD
 OS None
 OC Artificial sequences.
 PN JP 199705885 A/7
 PD 11-MAR-1997
 PP 29-MAR-1996 JP 1996099684
 PR 31 MAR 1995 JP 95P 9980, 19-JUN-1995 JP 95P 175540 PI
 KOSHIDA SHOGO, OKA YUMIKO, HIRAI YONEI
 PC C12N15/09, C07H21/04, C07K7/06, C07K14/485, PC
 C12P21/02//A61K38/00,
 PC. (C12P21/02/C12R1/19);
 CC strandedness: Double;
 CC topology: Linear;
 FH Key
 FH Location/Qualifiers
 FT source 1. .486
 FT /organism='Artificial sequences' FT
 FT mat_peptide 1. .483
 FT /product='Epimorphin fragment designated' FT
 FT /note='fragment (23) is derived from 104-264
 FT aa of mouse
 FT Epimorphin'.
 FEATURES source
 /organism="unidentified"
 /ab_xref="taxon:32644"
 /db_xref="taxon:32644"
 BASE COUNT 152 a 110 c 139 g 85 t.
 ORIGIN

Query Match 9.6%; Score 124.8; DB 6; length 486;
 Best Local Similarity 60.7%; Pred. No. 3.9e-10; Mismatches 204; Conservative 0; Indels 0; Gaps 0;

Qy 1 atccaggccgcgtggatcatggccaaaggaaatctcgccggcggccatgtggatcg 60
 Db 133 ATCCAGCGCAGCTGGAGATCACTGGAGGACCAACTGAGGGATG 192
 Qy 61 ttccatggatggatggggactgtgtttccggaaacttgcgtggccacgtgaaaggcg 120
 Db 193 CTGGAGAGCGGGAGGCCATCTCATCTGGATATTATTCAGTTCAAACT 252
 Qy 121 cggcgcgcctcaacagatcgagacgcggccaaactgtggcgctggagggcg 180
 Db 253 AGGCAAAGCTCTCATGAGTCAGTCAGGCCAAAGACATCATGAAGCTGGAGACCG 312
 Qy 61 ttccatggatggatggggactgtgtttccggaaacttgcgtggccacgtgaaaggcg 120
 Db 193 CTGGAGAGCGGGAGGCCATCTCATCTGGATATTATTCAGTTCAAACT 252
 Qy 121 cggcgcgcctcaacagatcgagacgcggccaaactgtggcgctggagggcg 180
 Db 253 AGGCAAAGCTCTCATGAGTCAGTCAGGCCAAAGACATCATGAAGCTGGAGACCG 312
 Qy 181 atccggacatcaacagacttttgcataatggcggtgtggatggaaaggccgc 240
 Db 313 ATCCAGCGCAGCTGGAGATCACTGGAGGACCAACTGAGGGATG 372
 Qy 241 accctgaacgtcatcgactcaacggtacaaggcggtcgactacccggcaggc 300
 Db 373 ATGGTCAACAACATTCAGAGAACGACGTGGTACCTCGTAGATACGGAGACATGGCAAG 432

				Gencore version 4.5
Copyright (c) 1993 - 2000	Compugen Ltd.			
On nucleic - nucleic search, using sw model				
Run on:	September 24, 2002, 14:13:36 ; Search time 237.5 seconds			
	(without alignments)			
Searched:	US-09-762-249-13			
Perfect score:	1295			
Sequence:	1 atccageggcaggat.....aaaaaaaaaaaaaaa 1295			
Scoring table:	IDENTITY.NUC			
	Gapop 10.0 , Gapext 1.0			
Minimum DB seq length:	0			
Maximum DB seq length:	200000000			
Post-processing:	Minimum Match 0%			
	Maximum Match 100%			
Database :	Listing first 45 summaries			
N_Geneseq_032802:*				
1: /\$IDS1/gcgttca/geneseq/geneseq-emb1/NA1980.DAT:*				
2: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1981.DAT:*				
3: /\$IDS1/gcgata/geneseq/geneseq-emb1/NA1982.DAT:*				
4: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1983.DAT:*				
5: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1984.DAT:*				
6: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1985.DAT:*				
7: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1986.DAT:*				
8: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1987.DAT:*				
9: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1989.DAT:*				
10: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1990.DAT:*				
11: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1991.DAT:*				
12: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1992.DAT:*				
13: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1993.DAT:*				
14: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1994.DAT:*				
15: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1995.DAT:*				
16: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1996.DAT:*				
17: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1997.DAT:*				
18: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1998.DAT:*				
19: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA1999.DAT:*				
20: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA2000.DAT:*				
21: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA2001A.DAT:*				
22: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA2001B.DAT:*				
23: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA2002.DAT:*				
24: /\$IDS1/gcgatc/geneseq/geneseq-emb1/NA2002.DAT:*				
	SUMMARIES			
Result No.	Score	Query Match Length	DB ID	Description
1	209.2	16.2	240	21 AAZ61486
2	124.8	9.6	486	17 AAT16088
3	124.8	9.6	486	18 AAT6243
4	124.8	9.6	564	17 AAT16087
5	124.8	9.6	564	18 AAT6212
6	124.8	9.6	711	17 AAT16086
7	124.8	9.6	711	18 AAT62411
8	124.8	9.6	798	17 AAT16090
9	9.6	798	18 AAT62414	
				ALIGNMENTS
RESULT	1			
ID	AA261486			standard; cDNA; 240 BP.
XX				
AC	AA261486;			
XX				
DT	19-JUN-2000			(first entry)
XX				
DE				cDNA sequence of tumour suppressor activated pathway gene TSAP21.
XX				
KW				Tumour suppressor activated pathway gene; TSAP; apoptosis;
KW				tumour suppression; tumour suppressor inhibited pathway gene; TSIP3;
KW				cancer; antiviral; ss.
OS				Homo sapiens.
XX				
PN				FR2782085-A1.
XX				
PD	11-FEB-2000.			
XX				
PF	05-AUG-1998;			98FR-0010077.
XX				
PR	05-AUG-1998;			98FR-0010077.
XX				
PA	(DAUS-) FOND DAUSSET-CEPH JEAN.			
XX				
PT	Anson R, Telerman A;			
XX				
DR	WP1; 2000-208788/19.			
XX				
PT	Nucleic acid sequences useful for gene therapy of cancer and viral infections corresponds to genes whose expression is induced or inhibited			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

XX
PN JP09065885-A.
XX
PD 11-MAR-1997.
XX
PF 29-MAR-1996; 96JP-0099684.
XX
PR 19-JUN-1995; 95JP-0175540.
XX
PR 31-MAR-1995; 95JP-0099980.
XX
(SUME) SUMITOMO ELECTRIC IND CO.
XX
DR WPI; 1997-220419/20.
XX
DR P-PSDB; AAW14262.
XX
The invention relates to novel human (AAW14257-9) or mouse (AAW14260-2) epimorphin proteins with replacements, deletions or substitutions in the amino acid sequence. The new epimorphin protein consists of: (a) an N-terminal coiled-coil region; (b) a functional domain in the middle; and (c) a C-terminal coiled-coil region. A hydrophobic region in the C-terminal has been deleted and at least some amino acids have been deleted from the terminals of coiled coil regions (a) and/or (c). This sequence encodes the mouse epimorphin protein lacking the N-terminal 103 amino acids, i.e. the complete N-terminal coiled-coil domain.
CC Epimorphin is a protein which regulates morphogenesis of epithelial tissue; It can be used for treatment of tissues or used directly in artificial organs or as an ingredient in cosmetics, hair tonic, etc.
CC Sequence 486 BP; 152 A; 110 C; 139 G; 85 T; 0 other;
XX
Query Match 9.6%; Score 124.8; DB 18; Length 486;
Best Local Similarity 60.7%; Pred. No. 3_9e-13;
Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
Oy 1 atccacgcccgcgtggagatcatggcaaggagtctcgccggaccatcgaggacatg 60
Db 133 atccagcgccaggactggagatcatctggagaccaccactgacgacgactggaaagatg 192
Oy 61 ttccgaggaggtaatggggactgtttcccgagaaacttgcgtggccgacgtggggcg 120
Db 193 ctggagagcgggaaagccgtccatctccatcgatattatcatcagatttcacaaatcaact 252
Ov 121 cggccccccctcaacaggatcgaggccggccacactgtcgccctcgagaccgc 180
253 agcaagactctcaatagatcgatcgatccggccacaagacatcatgaaagctggagaccgc 312
Oy 181 atcccgacgtcacacggatcttcgtcgaggcggtgtggagaaaggccgcac 240
Db 313 atccggagatcgacgatgttcgtatggccatgttgtgtcgagacttaggtgaa 372
Oy 241 acctcgaaacgttgcgtggatcatggcaaggaaacgcgtcgacatccggccggcc 300
Db 373 atccgtcaacacatcgaggaaatgtggatctgtggatcatcgatggacatcgcaag 432
Oy 301 ggcgtgtggaaaggccgttcacgtggagaaag 336
Db 433 gaagagacgaagaaaggccatcaaataccaggaccaag 468
RESULT 4
ID AAT16087 standard; cDNA; 564 BP.
XX
AC AAT16087;
XX
DT 30-MAY-1996 (first entry)
XX
DE Mouse epimorphin fragment (3M).
XX
PN XX
KW Epimorphin; human; mouse; wound; burn; epithelial tissue;
KW diagnosis; treatment; morphogenetic abnormality; cosmetic;
KW hair growth stimulator; ds.
XX
OS XX
Mus musculus.
XX
EP698666-A2.
XX
PR 28-FEB-1996.
XX
PA PA (SUME) SUMITOMO ELECTRIC IND CO.
XX
PR 20-JUN-1995; 95EP-0304270.
XX
PI Hirai Y, Koshida S, Oka Y;
XX
PR 31-MAR-1995; 95JP-0099980.
XX
DR 21-JUN-1994; 94JP-0162874.
XX
PR 31-MAR-1995; 95JP-0099979.
XX
P-PSDB; AAR92047.
XX
PA PA (SUME) SUMITOMO ELECTRIC IND CO.
XX
PR 31-MAR-1995; 95EP-0099980.
XX
PT 21-JUN-1994; 94JP-0162874.
XX
DR P-PSDB; AAR92047.
XX
PT Novel polypeptide containing an epimorphin functional domain - has possible benefits in epithelial tissue treatments, e.g. burns and for artificial organs
XX
PS Claim 27: Page 31; 62pp; English.
CC New polypeptides contain a first portion of 5-99 amino acids joined to a second portion coning. at least a functional domain of epimorphin. The first portion may be selected from the peptides given in AAR92029 to AAR92036. The second portion may be full-length epimorphin (see AAR92037 to AAR92042 for human and mouse epimorphins).
CC Fragments (123), (2M), (3M) and (23) of epimorphin given in AAR16083 to AAT16090 are used in the prodn. of modified epimorphins.
CC 123: N-terminus to right before C-terminal hydrophobic domain.
CC 2M: amino acid 30 to right before C-terminal hydrophobic domain.
CC 3M: amino acid 79 to right before C-terminal hydrophobic domain.
CC 23: amino acid 105 to right before C-terminal hydrophobic domain.
CC The modified epimorphins are useful for the development of diagnosis and treatment of morphogenetic abnormalities of epithelial tissue or novel remedies for wounds, e.g. burns, after surgery and for artificial organs. They may also be used as components of cosmetics, hair growth stimulators, etc.
XX
Sequence 564 BP; 181 A; 124 C; 162 G; 97 T; 0 other;
XX
Query Match 9.6%; Score 124.8; DB 17; Length 564;
Best Local Similarity 60.7%; Pred. No. 4e-13;
Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
Oy 1 atccagcccgatgtggatcatggcaaggaaatgttcgtggccggaccatcgaggatg 60
Db 211 atccagcccgatgtggatcatctggggggaccaccactgtacgcacgtggaaagatg 270
Oy 61 ttccgaggaggtaatggggactgtttcccgagaaacttgcgtggccgacgtggggcg 120
Db 271 ctggagccggaaagccgtccatcttcctctcgatattatcatcagatttcacaaatcaact 330
Oy 121 cggccccccctcaacaggatcgaggccggccacactgtcgccctcgagaccgc 180
Db 331 aagcaagactctcaatagatcgatcgatccggccacaatcatcgaaactgtggagaccgc 390
Oy 181 atcccgacgtcacacggatcttcgtcgaggatggccgtgtggagaaaggccgcac 240
Db 391 atccggagatcgacgatgttcgtggatcatggataggccatgttgctcgagactcagggtgaa 450
Oy 241 acctcgaaacgttgcgtggatcatggcaaggaaacgcgtcgacatccggccggcc 300

OY	1 atccagcgccgactggaggatcatggcaaggagtcggggaccactgtacggaggatcg 60	CC	of epithelial tissues. It can be used for treatment of tissues or used
DB	358 atccagcgccgactggaggatcatggaggaccactgtacggaggatcg 417	CC	directly in artificial organs or as an ingredient in cosmetics, hair
OY	61 ttccagcgggttaatggggactgtttccggagaactgtctggccgaactgtggggccg 120	CC	tonic, etc.
DB	418 ctggaggcgccaaggcgtccatctcatcgatattatcatcgatcaaataact 477	XX	
OY	121 cggccccccctaaacggatccaggccggccggccactgtcgccgttgagagccg 180	Sequence	711 BP; 232 A; 155 C; 201 G; 123 T; 0 other;
DB	478 aggcaagctctcaatgtggatcgatcgtccggcccaaaagacatcatcgaaatcg 537	SQ	
OY	181 atccgcgacgtcacggcgtcttgcagaggcggtgtggggaaaggaggccg 240	Query Match	9.6%; score 124.8; DB 18; Length 711;
DB	538 atccgagagctcacggatgttcatggatatggccatgttcgagactcgaggcgaa 597	Best Local Similarity	60.7%; Pred. No. 4.2e-13; 0; Mismatches 132; Indels 0; Gaps 0;
OY	241 accctggaaatgtatcgaggctcaacgtacaaagacggctgactacacggccaggccaa 300	Db	358 atccagcgccgactggaggatcatggcaaggactctcgccgtggggccg 417
DB	598 atggccacaadatcgagagaatgtgttagatcgaggaaatcgccaaatcgccaa 657	Query	1 atccagcgccgactggaggatcatggcaaggactctcgccgtggggccg 60
OY	301 ggcgcaggcgccggaaaggccgtggatcgaggcgaaag 336	Db	61 ttccagcgggttaatggggactgtttccggagaactgtctggccgaactgtggggccg 120
DB	658 gaagagacgcaagaagccatcaaataccagagcaag 693	Db	121 cggccccccctaaacggatccaggccggccggccactgtcgccgttgagagccg 180
RESULT	7	Db	478 aggcaagctctcaatgtggatcgatcgtccggcccaaaagacatcatcgaaatcg 537
ID	AAT62411	QY	181 atccgcgacgtcacggcgtcttgcagaggcggtgtggggaaaggaggccg 240
XX	AAT62411 standards; cDNA; 711 BP.	Db	538 atccgagagctcacggatgttcatggatatggccatgttcgagactcgaggcgaa 597
AC	AAT62411;	QY	241 accctggaaatgtatcgaggctcaacgtacaaagacggctgactacacggccaggccaa 300
DR	02-JUL-1997 (first entry)	Db	598 atggccacaadatcgagagaatgtgttagatcgaggaaatcgccaaatcgccaa 657
XX	Mouse epimorphin truncated variant 2M coding sequence.	QY	301 ggcgcaggcgccggaaaggccgtggatcgaggcgaaag 336
XX	Human; mouse; epimorphin; coiled-coil region; functional domain; tissue; hydrophobic; deletion; truncation; regulation; morphogenesis; epithelium; artificial organ; cosmetic; hair tonic; ds.	Db	658 gaagagacgcaagaagccatcaaataccagagcaag 693
KW	Synthetic.	RESULT	8
OS		ID	AAT16090 standard; cDNA; 798 BP.
XX		XX	AAT16090;
PN	JP09065885-A.	AC	
PD	11-MAR-1997.	XX	
XX	29-MAR-1996; 96JP-0099684.	DT	06-JUN-1996 (first entry)
PR	19-JUN-1995; 95JP-0175540.	XX	Mouse epimorphin fragment (123).
PR	31-MAR-1995; 95JP-0099980.	XX	
XX	(SUME) SUMITOMO ELECTRIC IND CO.	KW	Epimorphin; human; mouse; wound; burn; epithelial tissue; diagnosis; treatment; morphogenetic abnormality; cosmetic; hair growth stimulator; ds.
PA		KW	
XX		OS	Mus musculus.
DR	WPI; 1997-220419/20.	XX	
DR	P-PSDB; 14260.	PN	EPP8666-A2.
PT	Modified epimorphin and related DNA - useful e.g. for treatment of tissues or in artificial organs, or as an ingredient in cosmetics	XX	
PT	Example 1; Page 13; 18pp; Japanese.	PD	28-FEB-1996.
PS		XX	
XX		PF	20-JUN-1995; 95EP-0304270.
XX		XX	
PR	31-MAR-1995; 95JP-0099980.	PR	31-MAR-1995; 94JP-0162874.
PR	21-JUN-1994; 94JP-0099979.	PR	31-MAR-1995; 95JP-0099979.
PR		XX	
XX		PA	Hirai Y, Kosida S, Oka Y;
CC	The invention relates to novel human (AAW14257-9) or mouse (AAW14260-2) epimorphin proteins with replacements, deletions or substitutions in the N-terminal coiled-coil region; (b) a functional domain in the middle; and (c) a C-terminal coiled-coil region. A hydrophobic region in the C-terminal has been deleted and at least some amino acids have been deleted from the terminals of coiled coil regions (a) and/or (c). This sequence encodes the mouse epimorphin protein lacking the N-terminal 29 amino acids. Epimorphin is a protein which regulates morphogenesis	DR	WPI; 1996-118213/13.
CC		XX	
CC		PT	Novel polypeptide containing an epimorphin functional domain - has
CC		PT	Possible benefits in epithelial tissue treatments, e.g. burns and
CC		PT	for artificial organs
CC		XX	Claim 27; Page 51; 62pp; English.

Db 565 agccaagcttcataatgagatcgaggccgcctaaagacatacatgaaatggaggaccaggc 624
 QY 181 atcccgagcacgtacacgactcttgcagatggcgctgtggaaaggccg 240
 Db 625 atccgagagactcacatcgatgttcataatggccatgttgtaatcgatggatgg 684
 QY 241 accctgaacctcgatcgactcaacgtacaacgacgtcgactacacggccaggcaag 300
 Db 685 atggtaacaacaaatcgagaatgtgtgactctgttagatcatcgatggaaatgc 744
 QY 301 gicgcagggtcgaaaggccgtcgatcgaggagaag 336
 Db 745 gaagagacgcaagaaaggccatcaaataccagacgcaag 780

RESULT 14
 ID AAQ41592
 XX AAQ41592 standard; cDNA; 870 BP.
 AC AAQ41592;
 XX
 DT 24-AUG-1993 (first entry)
 XX
 DE mouse epimorphin coding sequence.
 XX
 KW vascularisation; induction; epithelial tissue morphogenesis; ds.
 XX
 OS mus musculus.
 XX
 WO9308213-A.
 FH Key Location/Qualifiers
 FT CDS 1..870
 FT /*tag= a
 FT /product= epimorphin
 PD 29-APR-1993.
 XX
 PR 15-OCT-1992; 92W0-JP01340.
 XX
 PR 16-OCT-1991; 91JP-0294855.
 PR 16-OCT-1991; 91JP-0294857.
 PR 17-APR-1992; 92JP-0122906.
 PR 30-APR-1992; 92JP-0135692.
 XX
 PA (BIOM-) BIOMATERIAL RES INST CO LTD.
 XX
 PT Hirai Y, Takashina M, Takebe K;
 XX
 DR WPI; 1993-152423/18.
 DR P-PSDB; AAR36552.
 XX
 PT Novel active substance epimorphin, its gene and antibodies - for
 PT diagnosing and treating epithelial diseases - for
 PS Claim 10; Page 55; 76pp; Japanese.
 XX
 CC This sequence encodes one of three isolated isoforms of mouse
 CC epimorphin, a protein produced by mesenchymal cells and which
 CC induces epithelial tissue morphogenesis. Modified forms of
 CC epimorphin in which the C-terminal hydrophobic region is replaced
 CC by a defective or non hydrophobic peptide are also claimed.
 CC See also AAQ41593 and AAQ41594.
 XX
 Sequence 870 BP; 261 A; 192 C; 253 G; 164 T; 0 other;
 SQ

Query Match 9.6%; Score 124.8; DB 14; Length 870;
 Best Local Similarity 60.7%; Pred. No. 4.3e-13;
 Matches 204; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

Db 445 atccageccaggatggagatcaatggggaccactcgacgactggaaaggatg 504
 QY 61 ttccagcggatgtttccggaaactgtctggccggatgtggggcc 120
 Db 505 ctggagacccggaaatccatcttcataatcgatattatcgatccaaatcc 564
 QY 121 cggccgcctcaacgatcgaggccacccggactgtgcgcgtggagcc 180
 Db 555 agccaagatctcaatgatcgatcgatggccggccaaagccatcgatgg 624
 QY 181 atcccgagctacacgactttgtgatcgatggatgg 300
 Db 625 atccgagatgttcataatggccatgttgtaatcgatggatgg 684
 QY 241 accctgaacctcgatcgactcaacgtacaacgacgtcgactacacggccac 240
 Db 685 atggtaacaacaaatcgagaatgtgtgactctgttagatcatcgatgg 744
 QY 301 gicgcagggtcgaaaggccgtcgatcgaggagaag 336
 Db 745 gaagagacgcaagaaaggccatcaaataccagacgcaag 780

RESULT 15
 ID AAQ75247
 XX AAQ75247 standard; cDNA; 870 BP.
 AC AAQ75247;
 XX
 DT 10-AUG-1995 (first entry)
 XX
 DE wild type mouse epimorphine gene.
 XX
 KW Probe; epimorphine; human; mouse; lambda-gtl; expression library;
 KW monoclonal antibody; isoform; drug; congenital; acquired; E.coli;
 KW epidermal abnormality; ds.
 XX
 OS Mus musculus.
 XX
 PN JP06293800-A.
 XX
 PD 21-OCT-1994.
 XX
 PR 15-OCT-1992; 92JP-0301581.
 XX
 PR 15-OCT-1992; 92JP-0301581.
 XX
 PA (BIOM-) BIOMATERIAL KENKYUSHO KK.
 XX
 DR WPI; 1995-00638/02.
 DR P-PSDB; AAR6479.

PT Human or murine epimorphine - useful for development of drugs to
 PT treat congenital and acquired epidermal form abnormality
 PT
 PS Claim 10; Page 8; 41pp; Japanese.
 XX
 CC The sequence of the wild type gene encoding mouse epimorphine. A DNA
 CC fragment (AAQ75250) containing the mouse epimorphine gene was isolated
 CC from a lambda-gtl expression cDNA library screened with a monoclonal
 CC antibody raised against mouse epimorphine. A probe (AAQ75243) derived
 CC from the mouse gene sequence was used to isolate isoforms of the mouse
 CC gene (AAQ75248-9) and the gene encoding human epimorphine (AAQ75244) and
 CC isoforms (AAQ75245-6). The genes were cloned into expression systems for
 CC the production of the protein in E.coli and in animal cells. The
 CC epimorphine can be used in the development of drugs to treat both
 CC congenital and acquired epidermal form abnormality.
 XX
 Sequence 870 BP; 261 A; 192 C; 253 G; 164 T; 0 other;

Query Match 9.6%; Score 124.8; DB 16; Length 870;

Db	505	CTGGAGGGGAAAGCGTCCATCTCATCTGGATAATTATCAGATTCACAATCACT	564
Qy	121	cggccgcctcaacagatcgagggccacccggcgaaactgtcgccggcgaggcgcc	180
Db	565	AGCAAGCTCTCAATCGATCGAGTCCGCCAACAAAGACATCATGAGCTGGAGACCCAGC	624
Qy	181	atccggcagtgacacgagcttcttcgcaatggcggtgtcggtggagaaaggcagccgac	240
Db	625	ATCCGAGACGTCACCGAGATGTGTCAGTATGGCATGGCAGTGTGTCGAGACTCAGGGTGAA	684
Qy	241	accctaactgtcatcgagtcacacgtttacaatggcggtcaactacaccggccaggccaa	300
Db	685	ATGGTCACAACTCGAGAATGTTGGACTCTGTAGATTACGGACATGCCAAC	744
Qy	301	gcgcagggtcgaaaggcggtcgatcgaggagaag	336
Db	745	GAAGAGACGAGAAAGCCATCAATACCGAGCAAG	780

RESULT 4
US-08-493-071-14
; Sequence 14, Application US/08493071
; Patent No. 6127149
GENERAL INFORMATION:
APPLICANT: Hirai, Yohei
APPLICANT: Koshiba, Shogo
APPLICANT: Oka, Yumiko
TITLE OF INVENTION: MODIFIED EPIMORPHIN
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSE: LOWE, PRICE, LEBLANC & BECKER
STREET: 99 CANAL CENTER PLAZA, SUITE 300
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/493,071
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Price, Robert L.
REGISTRATION NUMBER: 22,685
REFERENCE/DOCKET NUMBER: 715-107
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-684-1111
TELEFAX: 703-684-1124
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 798 base Pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-493-071-14

Db	505	CTGGAGGGGAAAGCGTCCATCTCATCTGGATAATTATCAGATTCACAATCACT	564
Qy	121	cggccgcctcaacagatcgagggccacccggcgaaactgtcgccggcgaggcgcc	180
Db	565	AGCAAGCTCTCAATCGATCGAGTCCGCCAACAAAGACATCATGAGCTGGAGACCCAGC	624
Qy	181	atccggcagtgacacgagcttcttcgcaatggcggtgtcggtggagaaaggcagccgac	240
Db	625	ATCCGAGACGTCACCGAGATGTGTCAGTATGGCATGGCAGTGTGTCGAGACTCAGGGTGAA	684
Qy	241	accctaactgtcatcgagtcacacgtttacaatggcggtcaactacaccggccaggccaa	300
Db	685	ATGGTCACAACTCGAGAATGTTGGACTCTGTAGATTACGGACATGCCAAC	744
Qy	301	gcgcagggtcgaaaggcggtcgatcgaggagaag	336
Db	745	GAAGAGACGAGAAAGCCATCAATACCGAGCAAG	780

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 12:53:01 ; Search time 1809.41 Seconds
(without alignments)
9659.810 Million cell updates/sec

Title:	US-09-762-249-13
Scoring table:	IDENTITY_NUC
Gapop:	10.0 , Gapext 1.0
Searched:	13736207 seqs, 678477542 residues
Total number of hits satisfying chosen parameters:	27472414
Minimum DB seq length:	0
Maximum DB seq length:	200000000
Post-processing:	Minimum Match 0%
Maximum Match 100%	
Listing first 45 summaries	
Database :	EST:*
1:	em_estba:*
2:	em_eatbun:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estpl:*
6:	em_estro:*
7:	em_htc:*
8:	em_htc:*
9:	gb.est1:*
10:	gb.est2:*
11:	gb.htc:*
12:	gb.gss:*
13:	em_gss_hum:*
14:	em_gss_inv:*
15:	em_gss_pn:*
16:	em_gss_virt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	728	562	739	10 BI518995
C. 2	658.6	50.9	911	10 BI519330
C. 3	543.2	41.9	571	10 BM14927
C. 4	521.0	40.2	598	10 BF892936
C. 5	487.4	37.6	525	10 BF894780
C. 6	465.4	35.9	525	10 BF892927
C. 7	456.4	35.9	480	10 BI158564
C. 8	447.8	34.6	470	9 AJ275149
C. 9	439.8	34.0	464	10 BF893339
C. 10	429.6	34.2	9 AW026277	
C. 11	421.4	32.5	434	9 AA227632
C. 12	416.4	32.2	499	9 AW500047
C. 13	403.4	31.1	429	9 BF893543
C. 14	402.4	31.1	426	9 AA262151
C. 15	397.2	30.7	2219	11 AK017897
C. 16	393.4	30.4	406	9 AW662909
C. 17	384.9	29.7	409	9 AW500048

ALIGNMENTS

RESULT	1
BI518995	BI518995 NIH_MGC_118 Homo sapiens mRNA clone IMAGE:511074 5', mRNA sequence.
DEFINITION	603061815F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:511074 5', mRNA sequence.
ACCESSION	BI518995
VERSION	BI518995.1
KEYWORDS	GI:15343687 EST
SOURCE	human
ORGANISM	Homo sapiens
REFERENCE	Mammali; Eutheria; Primates; Catarrhini; Hominidae; Homo. (bases 1 to 739)
AUTHORS	NIH-MGC http://mgc.ncbi.nih.gov/
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: gapbs-r@mail.nih.gov
Tissue	Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.	
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)	
DNA Sequencing By: Incyte Genomics, Inc.	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
Plate:	LLNL1529 row: n column: 03
High quality sequence stop:	738.
FEATURES	Location/Qualifiers
source	1. .+39 /organism="Homo sapiens" /ab_xref="Taxon:9606" /clone="IMAGE:521074" /clone_id="NIH_MGC_118" /tissue_type="Leukocyte" /lab_host="DH10B" /Note="vector: PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range


```

/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94306
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/393,985
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Sholtz, Charles K.
/ REGISTRATION NUMBER: 38,615
/ REFERENCE/DOCKET NUMBER: 8600-0152
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2097, base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: unknown
/ MOLECULE TYPE: cDNA to mRNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: Rat syntaxin 1A 3' end (encoding amino acids 4-288; GenBank M95734)
/ INDIVIDUAL ISOLATE:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 2..859
/ US-08-393-985-1

```

Query	Match	Score	Length
	Best Local Similarity	DB 1;	2097;
	Matches 206; Conservative 0;	Pred. No. 4-17;	
	Mismatches 148;	Indels 0;	Gaps 0;
Qy	1 atccagcgccgtggagatcatggaaaggaaactgtctggggaccaggatcgaggacatg 60		
Db	437 ATCCAGGGCAGCTGGATCATGGCGGACACGCCACTGGCTGGATCATGGACTCCAGATCG 496		
Qy	61 ttcgagcggtaagtgggacgttttcgagaactgtctggcgacytgaggcg 120		
nb	497 CTGGAGCTGGAAATCCGCCATCTTGCCTGGATCATGGACTCCAGATCG 556		
	121 cggccgcctcaacaggatcggacgttttccggatggggatctgtccggccggccgc 180		
Db	557 AAGCAGGCCCTCAGTGATCATGGATCATGGCCATGTGGAGGCCAGGGAG 616		
Qy	181 atccggcgttacgggttttccggatggggatctgtccggccggccgc 240		
Db	617 ATCCGGAGCTCACGGATATGGTCATGGACATGGTGGACTACGTGGAGGGCGTG 736		
Qy	241 accctgtggaaaggcgatcgaggtaactggggatggggccaggccaaag 300		
Db	677 ATGATGACAGGATCGATCATGGTGGACTACGTGGAGGGCGTG 736		
Qy	301 ggcgggtggaaaggcgatcgaggtaactggggatggggccaggccaaag 354		
Db	737 TCTGACACCAAGAACGGCTCAAGTACCGAAGCCAGGAAGAGATC 790		

PUBMED REFERENCE	11076861 4 (sites)	BASE COUNT	562 a /note="putative"
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	5	572 g
JOURNAL	Nature 409, 685-690 (2001)	557 t	1 others
REFERENCE	5 (bases 1 to 2219)		
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Built, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagai, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Horii, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sano, H., Sasaki, D., Schirml, L., Shibusawa, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Teijima, Y., Toda, T., Yamamoto, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gc.riken.go.jp, Fax: 81-45-503-9216)		
COMMENT	Please visit our web site (http://genome.gsc.riken.jp/) for further details.		
FEATURES	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand DNA was primed with primer [5'-GAGAGAGAAGGATCCAGTCAAGAGCTCTTTTTTTTTTTT-3'], cDNA was prepared by using trehalose-thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-KLRSAKTAATTAATAATTAAATTAATACTTCAGTAAATGCTT-3']. cDNA was cleaved with BamHI and Xhol. Vector: a modified pBluescript KS(+)- after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.		
SOURCE	Location/Qualifiers 1..2219		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="NCBI:1907940"		
	/clone_id="5810405C08"		
	/tissue_type="thymus"		
	/sex="male"		
	/clone_lib="RIKEN full-length enriched mouse cDNA library"		
	/dev_stage="adult"		
CDS	/note="data source:SPTR, source key:075558, evidence:ISS putative homolog to SYNTAXIN 11		
	/codon_start=1		
	/protein_id="BAB30994.1"		
	/db_xref="GI:12857383"		
	/translation="MKEDRRLAEQELQELSRSYDOOPDGDNDFDAPREDIVFENDLLESLYRVIQDIDENQLLIDYVRQLGRQNVRVLETSMRLESSKRTDNTSIKAKTKRGEGHQIQRQLOBLMGKMSGEOLDEMPFEGOKWMDVSFLNLADEPQKARALNEFESKURERLIRLEGRIRDYHFLQMLVVKQEDTLNVIYQKTLDYGEAKAQVKKKNPCTICCFRCPCPN"		
	2196..2201		
	/note="putative"		
	polyA_signal		
	polyA_site		